Aalto University
School of Business

A HEDONIC ANALYSIS OF FINLAND'S LARGEST CITIES

Henri Rujala

Bachelor's Thesis

Summer 2017

Economics

Thesis Supervisor: Pekka Ilmakunnas

ABSTRACT

The price of the dwelling is affected by its locational and structural characteristics which home buyers evaluate by personal preferences. This thesis uses data from six largest cities in Finland contributing with total 6538 observations. Using historical purchase data and a hedonic pricing model this bachelor thesis estimates values for different structural characteristics of dwellings. The aim of the study is to reveal how different structural characteristics are valued in different parts of Finland and create material for future discussion about Finnish real estate markets. First, I find that certain characteristics of the dwelling have value to the consumers. Second, I discovered that each of the six research cities has unique characteristics that affect to the action of consumers.

Keywords: Hedonic pricing model, Finnish housing markets, data-analysis

Table of Contents

I. INTRODUCTION

A home is a necessary good because everybody must live somewhere. Apartments vary in their size, quality, structural characteristics, and location. These characteristics are preferred with different ways and when consumers are purchasing a home they unveil their preferences via price that they are willing to pay for certain features of an apartment. A home with the higher density of desired features are sold at the higher price and otherwise. (Laakso & Loikkanen, 2004.)

Consumers can buy or sell their homes in real estate markets, where buyer and seller interact. A home is an exceptionally high valued good because the average home in Finland is worth over four times the average annual income (Laakso & Loikkanen, 2004). A home is a multidimensional good which market values is affected by consumers individual preferences as previously stated. Consumers also vary on their personal characteristics such as income, values and personal preferences. These factors make both demand and supply of homes heterogeneous. Heterogeneity makes real estate markets very thin because the amount of certain type of homes and active consumer in the market will vary. (Laakso & Loikkanen, 2004.)

High transaction costs are an essential part of the real estate markets. Transaction costs are generated from three sources which are searching expenses, transmission cost and taxing cost. Significant is also acknowledge the presence of psychological costs. The psychological costs are generated in the process of moving from one social environment to other. Another vital aspect of real estate markets is the information asymmetry between buyer and seller. The seller has much more information on the quality of the house than the buyer. These elements together account for low liquidity of real estate markets and consumers also rarely sell their homes. (Laakso & Loikkanen, 2004.)

This paper will proceed as follows: First, in section two we are going to have a comprehensive overview of existing literature and limitations of hedonic pricing model. Section three, will provide an extensive review of collected data and what issues are emerging due to the heterogeneity of the real estate markets. Section four, will provide an overview of the process of creating the hedonic pricing model and the results emerging from the model. Section 5, provides a conclusion about the final results from the model and discussion about future research.

II. RELATION TO EXISTING LITERATURE

Hedonic pricing model was first introduced by Automobile Manufacturers Association economist Andrew Court in the 1930s. Court created a foundation for important assumptions for a well specified model and he assumed that semi-log form of the model would be the best option to smoothen the model to more linear one. (Goodman, 1998.)

The hedonic pricing model was familiarized to housing markets by Polinsky & Rubinfeld (1977) to address the problem of valuing goods that are not directly traded in proper markets. Halvorsen and Pollakowski continued their research of real estate markets and recognized the significance of structural and locational characteristics. The functional form they used in their research was similar to corresponding research nowadays. (Halvorsen & Pollakowski, 1979.)

The hedonic pricing model is a model that uses observed prices to reveal the true market value of characteristics of differentiated products. The hedonic pricing model assumes that goods are only units that are linked to each other with varying characteristics and the observed prices are comparable to those observed prices of goods. The function to estimate implicit prices is to use econometric tools such as first stage regression. (Rosen, 1974.)

Essential is to understand that observed price characteristics cannot be used to estimate demand or supply of given good (Rosen, 1974). Also important is to acknowledge that Rosen (1974) assumed that consumers are making decisions in a perfect understanding and the price is formed on the market equilibrium. Goodman (1978) argued and represented critic against the equilibrium assumption by noting that housing markets are high of conversion cost, heterogeneity, and immobility which are violating the equilibrium assumption of Rosen.

Halvorsen and Pollakowski (1979) addressed the fundamental issue about the lack of theoretical foundation in hedonic pricing model to choose a functional form of the equation. The absence of solid theoretical foundation has caused that common practice has been to use the best form in a given situation. (Halvorsen & Pollakowski, 1979.)

The hedonic pricing model is a function of variables and the process of selecting variables may cause a problem. Shiller (2008) noted that there are an almost unlimited number of possible hedonic variables and this creates almost unlimited possible hedonic equations. The problem with almost unlimited possible combinations of hedonic variables is obvious because it is creating a ground for researchers to choose the variables that give wanted results. (Shiller, 2008.)

Economic literature and publications in hedonic pricing model in Finland have focused primarily in the capital area and Helsinki. (Bello 2009, Gustafsson &Wogenius 2014.) Prior publications on the topic found that location of home might be connected to household status and structural characteristics such as a number of rooms and age have unique effects to the price. (Bello, 2009.) Gustafsson and Wogenius (2014) confirmed the importance of floor area and they estimated that the floor area explained 81% of the price. Another essential notification is that preceding research in Finland has differentiated by functional form used in the research. Bello (2009) used a log-linear model with dependent variable logarithmic price and Gustafsson and Wogenius (2014) used linear-linear model.

III. DATA AND METHODOLOGY

The data for the research is gathered from Asuntojen.hintatiedot.fi service which is maintained by Ministry of Environment with the help of Housing Finance and Development Centre of Finland. The database collects all the realized transactions in the Finnish real estate markets in the past 12-months.

The data is provided to the service by the Central Federation of Finnish Real Estate Agencies which is an umbrella organization for associations and companies in the housing markets. The service makes possible to separate observations in the individual types of buildings such as townhouse, detached house and blocks of flat. Block of flat are the main target of this research and the only type of building downloaded from the data.

The collected data consist of total 6538 random sample observations from Finland six largest cities with total of 1 745 509 inhabitants. (Statistics Finland 1.) 2054 observations were from Helsinki, 693 from Espoo, 680 from Turku, 757 from Vantaa, 1349 from Tampere and 429 from Oulu. These observations are the final numbers of observations after deleting data with missing observations, double observations and large outliers. The downloaded data did consist of following variables:

City, District, Area, Price, €/m^2, Construction Year, Floor/Floors, Elevator and Condition. The following table describes the continuous variables of the collected data:

Table 1: Description of the variables in the cities

VARIABLES Medians	(1) Helsinki	(2) Espoo	(3) Turku	(4) Vantaa	(5) Tampere	(6) Oulu
m^2	57	**60**	59	53,5	53	55
Price	**229500**	201000	120000	155400	142000	100000
Construction Year	**1962**	1982	1971	1991	1986	1977
Observations	2,054	693	680	757	1,349	429

Description of the continuous variables underlines the dissimilarities of individual characteristics in the research cities. The surface area m^2 between cities is consistent with the data from the Finnish real estate markets as a whole and there is little variation between individual towns. (Statistics Finland 2.) We can observe that Helsinki have both the oldest and the most valuable dwellings. Espoo is having the biggest and Tampere the smallest dwellings when measuring with the floor size m^2. Dwellings are more modern in Vantaa measured with construction year and the dwellings are most inexpensive in Oulu.

Although there is observable dissimilarity in variables between cities, there is interest to examine the variation between individual districts. Next table describes variables m^2, Price and Construction Year between districts in the given data.

Table 2. Description of variables in the districts

VARIABLES Medians	(1) Varissuo Oulu	(2) Vuosaari Helsinki	(3) Kaleva Turku	(4) Värttö Oulu	(5) Armonkallio Turku	(6) Kaartinkaupunki Helsinki
m^2	76	63,5	45	65,5	56	72
Price	68000	195063,5	152950	140000	143000	489000
Construction Year	1978	1972	1961	1974	1964	1908
Observations	24	68	152	6	17	9

Description of continuous variables between districts reveals that each district has their own personal individualities. The price of dwelling, floor area and construction year are inconsistent between districts and we could assume that certain characteristics of a dwelling might be clustered inside districts in terms of price and construction year.

This phenomenon is named district segregation which means that people with same qualities will concentrate to the certain areas. The profound mechanism behind segregation is caused by consumer utility maximization meaning that consumer consider their income, social status and personal preferences when making purchase decisions. Consumer utility maximization together with heterogeneity of real estate markets causes nonrationality. Results is that people with same qualities focus to the certain areas. (Laakso & Loikkanen, 2004.)

Another considerable mechanism of segregation might be filtering. Filtering means that new blocks of flat are better in terms of structural quality which makes them attractive for wealthier consumers. Also, it’s reasonable to thing that filtering can work other way round with older and poor-quality blocks of flats. (Laakso & Loikkanen, 2004.)

When districts have individual fixed qualities it’s essential to control them in the research. The Fixed effects regression controls for omitted variables that vary across entities but are constant over time. (Stock & Watson, 2007.) When, there are district level qualities that vary across entities, we can use fixed effect-model as a control mechanism. The fixed effects regression makes possible to create individual constant for each district to minimalize district level fixed effects and omitted variable bias in the regression. (Stock & Watson, 2007.)

Fixed effects regression is a solution to control problems caused by district segregation. Example the construction year might be variable that is correlated with districts fixed effects. Qualities like location, districts social status and consumers personal preferences might have an effect on the price level of the district. Due to these reasons, the fixed effects regression is controlling the potential effects of construction year in the model and making the model more reliable. (Stock & Watson, 2007.)

IV. RESULTS AND DISCUSSION

Before I started picking potential regressors to the hedonic regression I did analyse the dependent variable price and which functional form would be best for the situation. I made various tests where I plotted dependent variable with continuous variables in the data and the results are in appendix 1.

I found that due to high density of small swellings in the data, the best functional form for the given hedonic regression would be log-linear model where dependent variable is logarithmic price and regressors of the floor area is linear. The scatterplots for construction year indicate that construction year plotted against linear or logarithmic price is quadratic. This is problematic for the hedonic regression because there is no linear relationship between dependent variable price. I also compared the normal distribution with the price variable distribution and found that logarithmic price variable did have best distribution when comparing to the normal.

I started to create regression with the most significant regressor m^2, which and all other variables are described more specifically in figure 1. I added regressors such as DRooms1, DRooms2 and Drooms3 which are describing if the dwelling is studio, two-room or three-room dwelling. I generated dummy variables ConstructionYear1, ConstructionYear2 and ConstructionYear3 which indicate the construction year of the dwelling. The variables Droom4 and ConstructionYear1 were ejected from the regression to avoid perfect multicollinearity. These four variables are the base variables of my hedonic regression model because they describe the blocks of flat where the dwelling is located and the structural characteristics of the dwelling.

After adding the base regressors I added variables that are more specific and that describe detailed structural specifications of the blocks of flat and the dwelling. The variables DLowest and DHighest describe if the dwelling is in lowest or in highest floor of the blocks of flat. These variables and their potential significance to the value of the dwelling is interesting because commonly people prefer higher floors.

DGood and DAverage are also more specific variables because they are variables that describe general condition of the dwelling and if it is in good or average condition. These variables are important in the regression because they indicate how consumer valuate condition of the dwelling in the given city. The condition of the dwelling and how much people value good condition might be linked to the general condition of the dwellings in the city.

The final two variables of the regression are also one of the most interesting. I created DElevatorHighest variable that states if the dwelling is located in the highest floor and if there is elevator in the blocks of flat. This variable creates interesting comparison with cities where are old blocks of flats without an elevator and cities with more modern blocks of flats. The Floors variable describes number of floors in the buildings and works as a control variable in the regression.

My method for the hedonic regression is simple. After creating relevant hedonic pricing model with selected variables, I will no longer change it between regression when changing city. I believe that using same variables in all regressions will reduce omitted variable bias and the bias with too many irrelevant variables in the regression. (Shiller, 2008.)

The final equation of Hedonic pricing model is the following:

$$\begin{aligned} LnPrice = \beta 0 + \beta 1 m^2 + \beta 2 ConstructionYear2 + \beta 3 ConstructionYear3 + \beta 4 DRooms1 \\ + \beta 5 DRooms2 + \beta 6 DRooms3 + \beta 7 DLowest + \beta 8 DHighest + \beta 9 DElevator \\ + \beta 10 DGood + \beta 11 DAverage + \beta 12 DElevatorHighest + \beta 13 Floors + \varepsilon \end{aligned}$$

The results of the final regression are described in the following Table 3. When reading table 3 is important to remember the functional form of the hedonic pricing model. The dependent variable is natural logarithm of the selling price and this causes that coefficients of the regression should be understood as a percentage changes in the price when a unit change in the measured variable. The results are described in the figure 2.

As we can see from the results in the figure 2, there is several variables that are statistically significant. Notable is that the significant variables differ between measured cities and almost every city has their own individual significant variables. The r-squared values of the model are between (0,6099-0,8670) in overall and indicating that the functional form used in the hedonic pricing model was proper solution. In the comparison, Bello (2009) created hedonic pricing model for Helsinki residential property market and got r-squared value of 0,728 with log-linear model. Also, Gustafsson and Wogenius (2014) got r-squared value of 0,914 with linear-linear model.

When observing the six base variables of the hedonic regression (m^2, DConstructionYear2, DConstructionYear3, Drooms1, DRooms2 and DRooms3) we notice that only variables m^2 and

ConstructionYear3 have expected significance level of one percent in all the cities. Variable m^2 coefficient varies between 0.0120-0,00562 and is highest in Turku and lowest in Vantaa. The coefficient of 0,0120 in Turku indicates that unit change of m^2 will add about 1,20% to the value of the dwelling. The variable ConstructionYear3 varies between 0,577-0,160 and is highest in Vantaa but lowest in Helsinki. The coefficient of 0,577 for Vantaa means that if the dwelling is constructed in the 2000s, consumers are willing to pay 57,7% more than dwelling that's construction year is smaller than 1960. The coefficient of 0,160 of Helsinki suggest that if the dwellings construction year is higher than 1999, consumers are willing to pay 16% more when comparing dwellings which construction year is smaller than 1960.

It is interesting to observe that there is observable link between general age of the building population and how age of the blocks of flat affect to the selling price. Helsinki were the building population is the oldest in terms of construction year, the coefficient for variable ConstructionYear3 is lowest of the observed cities. This observation could be explained by the quadratic relationship between construction year and price. In Helsinki, the oldest and the newest dwellings are the most valuable ones.

The high coefficient for variable ConstrtuctionYear3 in the Vantaa could be explained by the highest median construction year among cities. General building population in Vantaa is consisting of blocks of flats that are more modern in terms of construction year, so dwellings with construction year smaller than 1960 are minority in Vantaa. Consumer might also prefer more modern and high-quality dwellings over older one's due to the quality of the dwelling and that might cause the high coefficient for variable ConstructionYear3 in Vantaa.

The variable DConstructionYear2 varies in significance and sign between cities. In Helsinki, Turku and Vantaa the variable is significant at one percent level. Tampere and Oulu have significance level of five percent. The sign is negative in Helsinki, Turku, Tampere and Oulu. In Vantaa, the sign for the variable ConsttructionYear2 was positive. The coefficients of the variable DConstructionYear2 vary between minus 0,0857-0,0320 and its highest in Helsinki but lowest in Tampere. The coefficient of -0,0857 in Helsinki means that consumers are willing to pay 8,57% less from dwelling which construction year is between 1960-1999 when comparing dwellings with construction year smaller than 1960. In Vantaa, the variable DConstructionYear2 has positive sign and coefficient of 0,16 which means that construction year between 1960 and 1999 increases the value of the dwelling by 16% when comparing dwellings with construction year under 1960.

Helsinki and Turku have the two oldest population of blocks of flat in terms of construction year and Vantaa have the newest one. This could explain the one percent significance level in these cities. In Helsinki and Turku, the most desired dwellings by consumers seems to be the oldest and the newest ones. In Vantaa, the results could be explained by generally new population of blocks of flats. Either the way, the results generally support the discovery of the quadratic relationship between price and construction year.

The variables Drooms1, DRooms2 and Drooms3 vary in significance level and expected sign between cities. The variable DRooms1 have significance level of one percent level in Helsinki, Vantaa, Tampere and Oulu. In Turku, the variable DRooms1 was significant only in five percent level. Interestingly, sign was negative in Helsinki, Vantaa, Tampere and Oulu but positive in Turku. The negative coefficients vary between 0,297-0,124 and the highest coefficient of -0,297 were from Oulu and lowest coefficient of -0,124 from Helsinki. The coefficient of -0,297 from Oulu suggest that if the dwelling is studio and located in Oulu, consumers are paying 29,7% less when compared to four room dwelling. The positive coefficient of 0,113 in Turku suggest that if the dwelling is a studio and located in Turku, consumers are willing to pay 11,3% more when compared to four room dwelling.

The variable Drooms2 have significance level of one percent in Vantaa, Tampere and Oulu. In Turku, the variable has significance level of ten percent. The sign was negative in Vantaa, Tampere and Oulu but positive in Turku. The negative coefficients vary between 0,221-0,0823 and the highest coefficient of -0,221 were from Oulu and the lowest coefficient of -0,0823 were from Tampere. The coefficient of -0,221 from Oulu suggest that if the dwelling have two rooms and is located in Oulu, there will be 22,1% decrease in the total value when comparing to four room dwelling. The positive coefficient of 0,0816 in Turku suggest that if the dwellings is located in Turku and has two rooms, there will be 8,16% increase in the total value when comparing to four room dwelling.

The variable DRoooms3 have significance level of five percent in Helsinki, Vantaa and Turku. Espoo and Tampere have significance level on ten percent. The sign was positive in Helsinki and Turku but negative in Vantaa. The positive coefficients vary between 0,0702-0,0274 and the value of 0,0702 is from Turku and the value of 0,0274 is from Helsinki. The coefficient of 0,0702 suggest that if the dwelling have three rooms and is located in Turku, there will be 7,02% increase in the total value when comparing to four room dwelling. The negative coefficient of 0,044 in Vantaa suggest that if the dwellings is located in Vantaa and have three rooms, consumer are willing to pay 4,4% less when comparing to four room dwelling.

The variable DLowest have significance level of one percent in Helsinki and Vantaa. In Oulu, the significance level of the variable DLowest is five percent. As expected, the sign was negative in all the cities were the variable DLowest were significant. The negative coefficients vary between 0,0597-0,0308 and the highest coefficient were from Oulu and the lowest from Helsinki. The coefficient of -0,0597 from Oulu, suggest that if the dwelling is in the bottom floor and located in Oulu, there will be 5,97% decrease in the total value.

Interestingly, the variable DHighest is significant only in five percent level in Espoo but not in other cities. As expected, the sign is positive and the coefficient of 0,0796 from Espoo suggest that if the dwelling is in highest floor and located in Espoo, there will be 7,96% increase in the total value.

The variable DElevator is significant in one percent level in Oulu but not in other cities. As expected, the sign is positive and the coefficient of 0,0911 from Oulu suggest that if the dwelling is located in Oulu and the blocks of flat have elevator, the will be 9,11% increase in the total value.

The variable DGood have significance level of one percent in Helsinki, Vantaa, Turku, Tampere and Oulu. The sign was also positive in cities with significance level of one percent. The positive coefficients vary between 0,455-0,186 and the highest coefficient of 0,455 is from Oulu and the lowest coefficient of 0,186 is from Helsinki. The coefficient of 0,455 from Oulu suggest that if the dwelling have good quality and is located in Oulu, there will be 45,5% increase in the total value. Interestingly, Helsinki have the lowest coefficient of 0,186 which suggest 18,6% increase in the total value if the dwellings is good quality and located in Helsinki.

The variable DAverage have significance level of one percent in Helsinki, Vantaa, Tampere and Oulu. In Turku, the significance level of the variable DAverage is five percent. The sign was positive in cities with significance level of one and five percent. The positive coefficients vary between 0,368-0,0874 and the highest coefficient of 0,368 is from Oulu and the lowest coefficient of 0,0874 from Helsinki. The coefficient of 0,368 from Oulu suggest that if the dwelling is in average condition and located in Oulu, there will be 36,8% increase in the total value. In Helsinki, the coefficient of 0,0874 suggest that if the apartment is in good condition, there will be 8,74% increase in the total value.

The variable DElevatorHighest have significance level of one percent in Helsinki and significance level of ten percent in Espoo. As expected, the sign is positive in Helsinki but interestingly the sign is negative in Espoo. The positive coefficient of 0,0529 from Helsinki suggest that if the dwelling is in top floor and there is elevator in the blocks of flat, there will be 5,29% increase in the total value. The negative coefficient of -0,0690 from Espoo suggest that if the dwelling is in the top floor and there is elevator in blocks of flat, there will be -6,90% decrease in the total value.

The final variable floors which describes the height of the blocks of flat, were an attempt to control possible omitted variable bias in qualities that tall blocks of flats have. The higher the blocks of flat, the higher is possibility to have an elevator. Important is to notice that consumers might prefer differently the bottom or top floors of high blocks of flats.

V. CONCLUSION

Creating hedonic pricing model for structural characteristics of dwellings is hard. The final model is affected by the individual assumption of researcher and process of learning. The lack of profound theoretical agreement between researchers about which functional form performs best under certain conditions and unique characteristics of the real estate markets make creating the hedonic pricing model hard and interesting process. (Halvorsen & Pollakowski, 1979.)

The hands-on process of creating hedonic pricing model gave surprising findings that affected the final form of the model. The linearity tests with log-log, log-linear and linear-linear versions of the hedonic pricing model helped me to choose the best functional form for the model. My choice for the final model was a log-linear model with dependent variable logarithmic price and linear m^2 as a regressor. I also noticed that there was a quadratic relationship between construction year and price, so I was forced to create three dummy variables that describe certain periods of time to avoid problems with nonlinearity.

Structural characteristics of the dwelling are not traded in the public markets but their value to consumers can be interpreted from premiums that consumers pay from dwellings with these structural qualities. This is done by collecting large amounts of data from real estate markets about realized transactions and creating a model that captures the effect of those specific characteristics to the price of the dwelling.

This bachelor thesis offers a conclusion that certain characteristics of the dwellings have value to the consumers. The values of the characteristics to the consumers differ between six largest cities in Finland. This thesis displayed thirteen different variables that describe various characteristics of the dwelling. These variables describe structural characteristics such as construction year and information about the elevator in the blocks of flat.

This thesis discovered that each of the six research cities has unique characteristics that affect to the action of consumers. Characteristics such as the general age of the dwelling population, general price of the dwelling population and general floor space of the dwelling population create each city real estate market unique. These characteristics affect how consumer prefer certain qualities in certain cities. An example in Helsinki, where are the oldest and the most expensive dwellings people prefer combinations of specific qualities like top-floor with elevator. Individually, elevator and top-floor dwelling aren't significant characteristics but together their unlike combination is valued by consumers. This unlikely combination is created by the fact that some of the oldest blocks of flats in Helsinki are missing an elevator. Consumers are willing to pay a premium for the dwellings that are in top-floor but with the possibility to the elevator.

The variables ConstructionYear2 and ConstuctionYear3 did reveal how the general age of the dwelling population will affect consumer preferences about construction year of the given dwelling. A good example is Vantaa where the median dwelling is constructed in 1991 which is 29-years higher than the Helsinki where are the oldest dwellings. This thesis revealed that consumer in Vantaa value construction year differently than other cities and this can be lead the general conditions of the Vantaa real estate markets. In Vantaa, the effect of the construction year is backward when compared the overall effect that construction year caused in other cities. Consumers are willing to pay a premium for houses which have construction year higher than 1960. The premium is 16% for dwellings with construction year between 1960-1999 and 57,7% premium for dwellings that are constructed in the 2000s.

An interesting result was also how consumer value condition of the apartment in different cities. You would think that there is a relationship between construction year and how people prefer condition of the apartment. I first thought that consumer in Helsinki and Turku which share the oldest building population in Finland should have consumers that value highly good or average condition over bad condition of the dwelling. I was surprised that consumers in Oulu value these characteristics more than consumers in Helsinki and Turku. If the dwelling is in good condition consumer are willing to pay 45,5% premium and 36,8% premium for dwelling with average condition compared to the dwelling with bad condition.

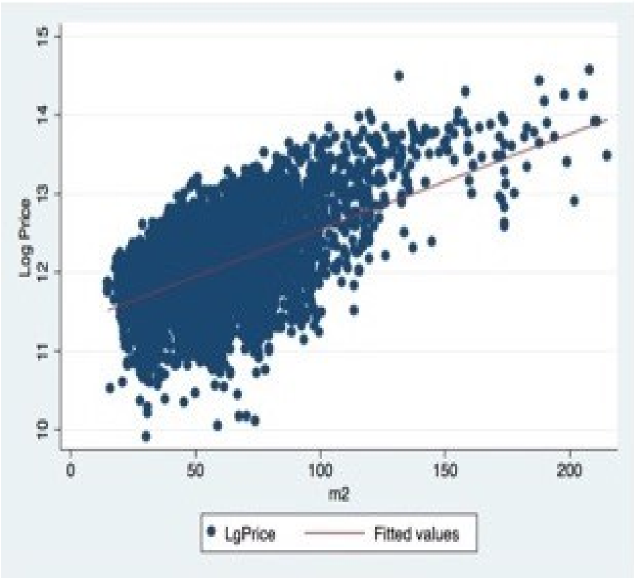
Another interesting result was that qualities such as elevator, bottom-floor and top-floor weren't widely preferred characteristics for consumers. Only consumers in Oulu pay a premium of 9,11% when a dwelling is located in the blocks of flat with elevator. Also, people in Espoo were only ones who pay a premium of 7,96% for a dwelling that is located in the top-floor. Surprising thing was that bottom floor was a significant character for consumers only in Helsinki, Vantaa and Oulu. Consumer pay 5,99% less in Oulu, 3,86% less in Vantaa and 3,08% less in Helsinki for a dwelling that is located in bottom-floor.

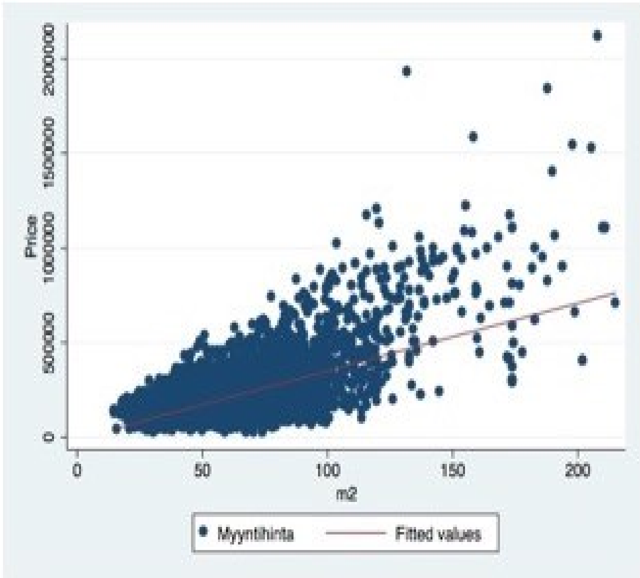
Finally, this bachelor thesis continued recent work on how researchers can use realized market transactions to estimate values of the dwellings. As most of the recent research and services of the private firms have focused on giving precise estimate of market value for dwelling. My research from this topic contributes to this discussion with giving details which characteristics of the dwelling are more valuable than other in the research cities. I hope that in the future my research can be used as a support function to the buying and selling decisions of individual consumers. As technology moves forward and consumer might be able to make markets transactions without leaving home, it becomes valuable to have tools that reveal which characters construct the price that automated service in future create for your dwelling.

Further research on the topic is needed, even though this thesis completes its purpose. It would be interesting to combine specific area information through GIS-data with structural information I used in this thesis. This would eliminate some of the biases caused by district segregation and it would create interesting research about how distance from city center and even metro stations effect to the selling price of the dwelling.

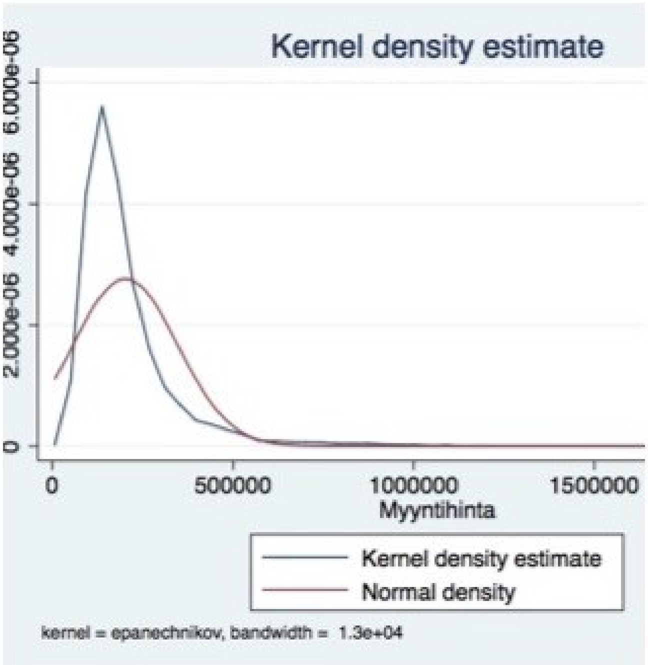
APPENDIX 1

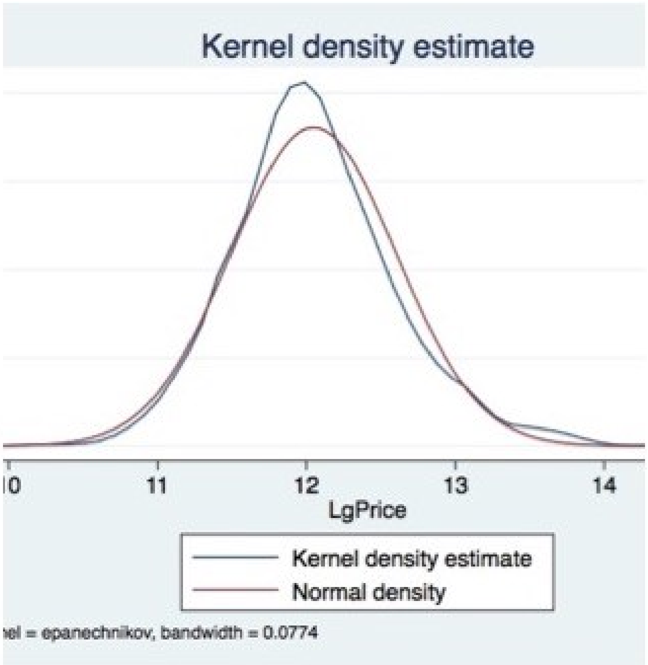
1.1 Linearity test. $\boldsymbol{m^2}$ against logarithmic and linear price





1.2 Kernel density estimate for logarithmic and linear price





1.3 Linearity test. Construction year against logarithmic and linear price

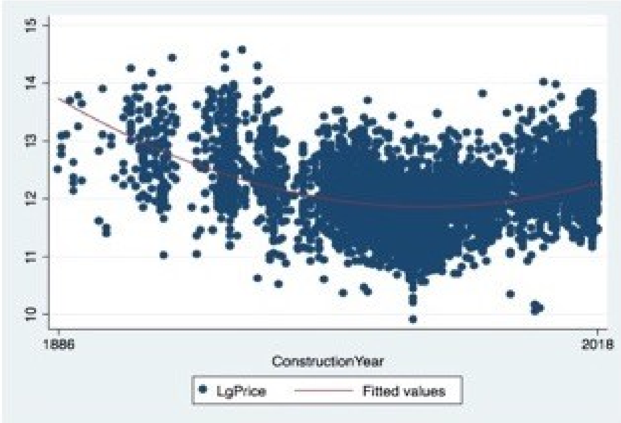


FIGURE 1: Description of variables

VARIABLE	DEFINITION
City	States which city the dwelling is located
District	States which District the dwelling is located
m^2	States the floor area of the given dwelling
ConstructionYear1	States if the construction year is smaller than 1960
ConstructionYear2	States if the construction year is between 1960-1999
ConstructionYear3	States if the construction year is higher than 1999
Floor	States the floor where the given dwelling is located
Floors	States the number of floors of the given blocks of flat
DElevator	States if there is elevator in the given blocks of flat
DHighgest	States if the dwelling is located at the top floor.
DGood	States if the dwelling is good condition
DAverage	States if the dwelling is in average condition
DBad	States if the dwelling is in bad condition
LgPrice	States the logarithmic price of the dwelling
Lg m^2	States the logarithmic area of the dwelling
DElevatorHighest	States if the dwelling is at top floor and building has elevator
DRooms1	States if the dwelling is studio
DRooms2	States if the dwelling has two rooms
DRooms3	Sates if the dwelling has three rooms
DRooms4	States if the dwelling has four or more rooms

FIGURE 2: Hedonic regression coefficients

VARIABLES	(1) Helsinki	(2) Espoo	(3) Turku	(4) Vantaa	(5) Tampere	(6) Oulu
m^2	0.0104***	0.0102***	0.0120***	0.00562***	0.0103***	0.00769***
	(0.000236)	(0.000894)	(0.000587)	(0.000648)	(0.000443)	(0.00115)
DConstructionYear2	-0.0857***	-0.00515	-0.0804***	0.160***	-0.0320**	-0.0738**
	(0.0101)	(0.0548)	(0.0234)	(0.0449)	(0.0148)	(0.0323)
DConstructionYear3	0.160***	0.380***	0.414***	0.577***	0.292***	0.360***
	(0.0139)	(0.0565)	(0.0312)	(0.0468)	(0.0159)	(0.0409)
DRooms1	-0.124***	0.0349	0.113**	-0.210***	-0.168***	-0.297***
	(0.0207)	(0.0632)	(0.0570)	(0.0418)	(0.0343)	(0.0857)
DRooms2	-0.0201	0.0343	0.0816*	-0.134***	-0.0823***	-0.221***
	(0.0162)	(0.0450)	(0.0438)	(0.0314)	(0.0266)	(0.0635)
DRooms3	0.0274**	0.0613*	0.0702**	-0.0440**	-0.0376*	-0.0713
	(0.0124)	(0.0322)	(0.0353)	(0.0222)	(0.0204)	(0.0464)
DLowest	-0.0308***	0.0124	-0.0270	-0.0386***	-0.0144	-0.0597**
	(0.00900)	(0.0208)	(0.0218)	(0.0122)	(0.0124)	(0.0251)
DHighest	-0.00558	0.0796**	0.0263	0.0173	0.0107	-0.00605
	(0.0122)	(0.0320)	(0.0350)	(0.0186)	(0.0192)	(0.0393)
DElevator	0.0141	0.00716	0.0310	-0.00681	-0.00706	0.0911***
	(0.00998)	(0.0187)	(0.0245)	(0.0165)	(0.0169)	(0.0321)
DGood	0.186***	-0.0405	0.204***	0.336***	0.228***	0.455***
	(0.0153)	(0.0516)	(0.0484)	(0.0305)	(0.0441)	(0.0785)
DAverage	0.0874***	-0.0641	0.0927*	0.227***	0.124***	0.368***
	(0.0153)	(0.0534)	(0.0487)	(0.0307)	(0.0449)	(0.0873)
DElevatorHighest	0.0529***	-0.0690*	0.00335	0.00280	0.0229	0.0402
	(0.0163)	(0.0393)	(0.0433)	(0.0220)	(0.0224)	(0.0473)
Floors	0.00203	0.0135***	0.0265***	0.00250	0.0110***	-0.00456
	(0.00248)	(0.00390)	(0.00541)	(0.00362)	(0.00353)	(0.00685)
Constant	11.60***	11.32***	10.52***	11.12***	10.99***	10.71***
	(0.0378)	(0.122)	(0.0978)	(0.0852)	(0.0729)	(0.142)
Observations	2,052	691	680	757	1,349	429
Number of DArea	124	58	81	47	83	65
District FE	YES	YES	YES	YES	YES	YES

Standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

REFERENCES

Bello, Omoniyi. 2009. Hedonic Analysis of Helsinki Residential Property Market. Master of Science Thesis. Royal Institute of Technology. Department of Real Estate and Construction Management.

Goodman, Allen. 1998. Andrew Court and the Invention of Hedonic Price Analysis. *Journal of Urban Economic*s 44: 291-298.

Goodman, Allen. 1978. Hedonic Prices, Price Indices and Housing Markets. *Journal of Urban Economics* 5, 471-484.

Gustafsson A, and Wogenius, S. 2014. Modelling Apartment Prices with the Multiple Linear Regression Model. Degree Project in Applied Mathematics and Industrial Economics. Royal Institute of Technology.

Halvorsen, Robert, and Henry O. Pollakowski. 1981. Choice of functional form for hedonic price equations. *Journal of Urban Economics* 10: 37-49.

Laakso, S, Loikkanen, H, A. 2004. *Kaupunkitalous. Johdatus kaupungistumiseen, kaupunkien maankäyttöön sekä yritysten ja kotitalousten sijoittumiseen.*

Polinsky, A.M., D.L. Rubinfeld. 1977. Property values and the benefits of environmental improvements: theory and measurement. In Public Economics and the Quality of Life. L. Wingo and A. Evans (eds). The Johns Hopkins University Press for Resources for the Future, Baltimore, Maryland.

Rosen, Sherwin. 1974. Hedonic Prices and Implicit Markets: Product Differentiation in Pure Competition. *The Journal of Political Economy* 82, no. 1: 34-55.

Shiller, J, Robert. 2008. Derivatives Markets for Home Prices. *National Bureau of Economic research. Working Paper 1396.*

Stock J. & Watson M. (2007) "*Introduction to Econometrics"*, Boston: Pearson/Addison Wesley, 2007.

Statistics Finland 1. Tilastokeskus. Suomen virallinen tilasto (SVT): Väestörakenne [verkkojulkaisu].ISSN=1797-5379. Helsinki: Tilastokeskus [viitattu: 31.7.2017]. Saantitapa: http://www.stat.fi/til/vaerak/index.html

Statistics Finland 2. Tilastokeskus. Suomen virallinen tilasto (SVT): Asunnot ja asuinolot [verkkojulkaisu].ISSN=1798-6745. 2016, Liitetaulukko 3. Pinta-ala huoneistoa kohti (m2) asunnon talotyypin mukaan 1970-2016, koko asuntokanta. Helsinki: Tilastokeskus [viitattu: 6.6.2017]. Saantitapa: http://www.stat.fi/til/asas/2016/asas_2016_2017-05-22_tau_003_fi.html